Supplementary information

Recruitment

Stroke subjects were prospectively recruited from the stroke service at Barnes-Jewish Hospital (BJH), with the help of the Washington University Cognitive Rehabilitation Research Group (CRRG)(Dr. Lisa Connor) from 5/1/2008 to 5/30/2013.

Inclusion criteria: (1) Clinical diagnosis of stroke at hospital discharge; (2) Persistent stroke symptom(s) at hospital discharge; (3) Awake, alert, and able to complete study tasks; (4) Age 18 or older.

Exclusion criteria: (1) Previous stroke, based on clinical imaging; (2) Multifocal stroke, based on clinical diagnosis or imaging; (3) >2 lacunes visible on head CT scans performed before study enrollment; (4) Schizophrenia, bipolar disorder, major depression, or other severe psychiatric condition; (5) Dementia (as noted in medical chart, as measured by a Short Blessed Score of 9 or greater, or as measured by a premorbid AD-8 score of 2 or greater); (6) Epilepsy, Parkinson's Disease, or other neurological disorder; (7) Brain injury; (8) End-stage renal disease, terminal cancer, class III or IV heart failure, or other diagnosis with a life expectancy less than 1 year; (9) Pre-morbid functional disability as measured by a Modified Rankin score of 2 or higher; (10) Claustrophobia; (11) Metal object in body precluding use of 3T MRI.

From the CRRG Registry, 6260 charts were screened. Of these, 5438 (87%) were excluded based on chart review alone: 1697 were discharged with a diagnosis other than stroke (such as complex migraine or TIA); of the remaining candidates, 1181 had a prior history of stroke; then 1188 had other neurological or psychiatric history listed as exclusion criteria; then 877 had other medical co-morbidities listed as exclusion criteria; then 387 had contraindications for MRI, and finally 108 showed multifocal lesions or excessive lacunae on clinical head CT. The 822 candidates screened in were all approached for enrollment.

Of these 822, 172 were enrolled in the study. The other 650 refused, were transferred to a facility outside the university hospital system, or remained medically unstable during the enrollment window (634), or were randomized to enrollment in a competing protocol (16).

Of the 172 patients enrolled, 40 were later excluded: 14 were unable to tolerate the scan; 2 had artifacts on the scan; 12 had tiny or questionable lesions; and 12 were found to have multifocal lesions or excessive lacunae.
See **Supplementary Figure 1** for Enrollment flowchart.

Healthy Control group

A healthy control group (N=30) was matched with the study sample for age, gender, and years of education. They serve as a performance baseline in the neuropsychological

and behavioral battery. They were enrolled using the same exclusion criteria as the stroke study group, from neighbors or spouses of stroke study participants.

Handedness

In the healthy control group 96% of subjects were right handed and 4% were left-handed. In the stroke group 91% were right handed and 9% were left handed.

Behavioral methods

The behavioral battery includes standard neuropsychological tests that have been extensively employed in stroke patients and have validated psychometric properties. The only computerized task (Posner) has been carefully characterized in terms of sensitivity and specificity(Rengachary et al., 2009). The behavioral battery includes mainly tests of body function according to the ICF classification. We include most of the tests recommended by NINDS and the Canadian Stroke Network task force for the harmonization of cognitive impairment measures in vascular diseases(Hachinski et al., 2006).

I. Motor Battery

Upper body function was measured in both arms as follows:

1. Active range of motion against gravity measured by goniometry at Shoulder Flexion, and Wrist Extension (Dreeben-Irimia, 2008);

2. Grip strength measured by dynamometry (Demeurisse et al., 1980);

3. Dexterity measured with the 9-Hole Peg Test in which patients placed nine plastic pegs into holes on a pegboard as quickly as possible (pegs/second) (Oxford Grice et al., 2003);

4. Function measured with the Action Research Arm Test total score (ARAT), in which patients performed functional grasp, grip, pinch, and gross motor movements according to the standardized protocol and were rated for quality of movement (van der Lee et al., 2001);

Lower body function was measured as follows:

1. Combined Walking Index: Patients were timed while walking 10 meters if able to safely do so unassisted. Patients who were unable to walk 10 meters were rated using the Walking item on the Functional Independence Measure. The following variable was recorded as a combined index of the two walking measures in order to capture variability both for maximally and minimally impaired patients: Score of 1: total assistance required to walk; Score of 2: maximal assistance required to walk; Score of 3: moderate assistance required to walk; Score of 4: minimal contact assistance required to walk; Score of 5: standby assistance required to walk; Score of 6: modified independence in walking (use of assistive device); Score of 7: independence in walking but a speed of <0.4 meters/second; Score of 8: independence in walking but a speed of 0.4 to 0.8 meters/second; Score of 9: independence in walking and a speed greater than 0.8 meters/second. Note that 0.4<meters/seconds<0.8 denotes household ambulation, while >0.8 meters/second denotes community ambulator (Kempen et al.,

2011) (Perry et al., 1995) (Keith et al., 1987);
2. Left/right Total Motricity Index (MI), which sums the manual muscle testing scores for left/right hip flexion, knee extension, ankle dorsiflexion; and
3. Ankle dorsiflexion goniometry for left/right active range of motion against gravity (Dreeben-Irimia, 2008).

II. Language Battery

Subtests of the Boston Diagnostic Aphasia Examination (BDAE-III) were performed according to the standard protocol (Goodglass et al., 2001):
1. Basic word discrimination: Patients pointed to the picture that matched the word named by the experimenter;
2. Commands: Patients performed one- to five-step commands spoken by the experimenter;
3. Complex Ideational Material: Patients answered yes/no questions spoken by the experimenter;
4. Boston Naming Test short form: Patients named the item pictured;
5. Oral Reading of Sentences: Patients read sentences aloud;
6. Comprehension of Oral Reading of Sentences: Patient answered multiple-choice comprehension questions about the sentences they just read;

In addition we measured phonetic/phonological processing with:

7. Non-word Reading: Four-letter nonwords (e.g. NORD) were presented. Patients were instructed to say the nonword aloud;
8. Stem Completion: Three-letter word stems (e.g., COU) were presented. Patients were instructed to say a word that started with those three letters (e.g., COUPLE). This task provides a very stereotypical pattern of language-related areas (Connor et al., 2006).

III. Executive Function

A subtest of the Delis-Kaplan Executive Function System was performed according to the standard protocol:
Animal Naming: Patients named as many animals as possible in 1 minute (Tombaugh et al., 1999) (Delis, 2001).

IV. Memory Battery

Visual memory was studied with the Brief Visuospatial Memory Test-Revised (BVMT-R) (Benedict, 1997). Patients studied abstract figures and were asked to reproduce them from memory on three immediate recall trials and one delayed recall trial. After the delayed recall trial, patients were shown figures and asked if each was one of the studied figures. Scores were calculated for the following variables:
1. BVMT Immediate Total Recall T-score: age-normed using the tables provided in the test manual;
2. BVMT Delayed Recall T-score: age-normed using the tables provided in the test manual;
3. BVMT Delayed Recall percent retained: calculated from the percent items retained

from last immediate recall trial to delayed recall; and
4. BVMT Delayed Recognition discrimination index: calculated from proportion of correct recognitions, correct rejections, misses, and false alarms, using the table provided in the test manual.

Verbal memory was assessed with the Hopkins Verbal Learning Test-Revised (HVLT-R)(Brandt and Benedict, 2001). Patients listened to a list of words and were asked to repeat them from memory on three immediate recall trials and one delayed recall trial. After the delayed recall trial, patients were read a list of words and asked if each was one of the studied words. Scores were calculated for the following variables:
1. HVLT Immediate Total Recall T-score: age-normed using the tables provided in the test manual;
2. HVLT Delayed Recall T-score: age-normed using the tables provided in the test manual;
3. HVLT Delayed Recall percent retained: calculated from the percent items retained from last immediate recall trial to delayed recall; and
4. HVLT Delayed Recognition discrimination index: calculated from the proportion of correct recognitions, correct rejections, misses, and false alarms, age-normed using the table provided in the test manual.

Spatial working memory was examined with a subtest of the Wechsler Memory Scale(Psychological Corp, 1981):
Spatial Span: Patients watched the examiner tap sequences on a block board and then asked to copy the sequences. After reaching a performance ceiling, the patients watched the examiner tap sequences on the block board and then asked to produce the sequences in reverse order. Scores were calculated for the following variables:
1. Spatial Span Forwards;
2. Spatial Span Backwards

V. Attention Battery

Different visuospatial attention processes were measured with the Posner orienting task (Posner et al., 1984). Stimuli were generated by an Apple Power Macintosh computer and displayed on a 17 inch Apple Monitor. Behavioral responses were acquired through a Carnegie Mellon button box interfaced with the computer. The experimenter visually screened for eye movements and encouraged visual fixation whenever a fixation break occurred. The display contained a central fixation cross and two eccentric, square frames (side 1 degree, center of frame at 3.3 degrees from the fixation cross) positioned along the horizontal meridian to the left and right of fixation. For the two patients with a quadrantanopsia, we presented stimuli in the visible part of the field on symmetrically opposite positions across the vertical meridian. The onset of a new trial was signaled by a color change, from red to green, of the fixation cross. Then 800 millisecond (ms) later an arrow cue pointing left or right appeared at fixation for 2360 ms. Following a delay ranging from 1000 to 2000 ms the target (an asterisk) appeared for 300 ms within one of the two frames (left or right). On 75% of the trials, the target appeared at the location indicated by the cue (valid condition), while on 25% of the trials it appeared at the opposite location (invalid condition). Patients had to detect the target as quickly as

possible with a key-press, using the ipsilesional hand. The RTs were recorded. An intertrial interval (ITI) of 2360 msec separated subsequent trials. Blocks contained 40 trials (30 valid, 10 invalid). Each patient completed 2 blocks. The test took a total of 15 minutes to administer, including a practice block. The following scores were calculated:

1. Visuospatial contralesional biases were measured with:
    a. *The Visual Field Reaction Times (RT):* relative delay in RTs for targets presented in the left vs. right visual field
    b. *The Visual Field Accuracy:* relative percent misses for targets presented in the left vs. right visual field
2. Deficits in shifting attention were measured with:
    a. *The Validity Effect RT*: relative delay in RTs for targets presented following valid vs. invalid cues
    b. *The Validity Effect Accuracy*: relative percent misses for targets presented following valid vs. invalid cues
3. Deficits in re-orienting to unattended locations with:
    a. *The Disengagement Effect RT:* relative delay in RTs for targets presented in the left visual field following an invalid cue.
    b. *The Disengagement Effect Accuracy:* relative percent misses for targets presented in the left visual field following an invalid cue.
4. Overall performance or sustained attention were measured with:
    a. Average RT: average of the RTs across all four conditions.
    b. Average accuracy: average percent misses across all four conditions.

Visuomotor spatial deficits were assessed with the Star Cancellation subtest of the Behavioral Inattention Test (BIT)(Wilson et al., 1987), and the Mesulam Unstructured Symbol Cancellation Test(Mesulam, 1985). The following score were calculated:

1. *Mesulam Center-of-cancellation, L-R misses:* which reflects the lateralized center of mass of misses, using the software provided by Rorden and Karnath, for left-sided vs. right-sided misses (Rorden and Karnath, 2010).
2. *BIT: star cancellation, Center-of-cancellation, L-R misses*: which reflects the lateralized center of mass of hits, using the software provided by Rorden and Karnath, for left-sided vs. right-sided misses.

Imaging methods

MRI Procedure and Scanning

Scanning was performed with a Siemens 3T Tim-Trio scanner at the School of Medicine of the Washington University in St. Louis including: structural, functional and diffusion tensor scans. Structural scans consisted of: (1) a sagittal MP-RAGE T1-weighted image (TR=1950 msec, TE=2.26 msec, flip angle=9 deg, voxel size=1.0 x 1.0 x 1.0 mm, slice thickness = 1.00 mm); (2) a transverse turbo spin-echo T2-weighted image (TR=2500 msec, TE=435 msec, voxel-size=1.0 x 1.0 x 1.0 mm, slice thickness = 1.00 mm); and (3) a sagittal FLAIR (fluid attenuated inversion recovery) (TR=7500 msec, TE=326 msec, voxel-size=1.5 x 1.5 x 1.5 mm, Slice thickness = 1.50 mm).

Data Preprocessing

Subject's T2-weighted and FLAIR images were co-registered with the T1-weighted MP-RAGE, in both cases using a cross-modal procedure based on alignment of image gradients (Rowland et al., 2005). The MP-RAGE was then transformed to an atlas-space (Talairach and Tournoux, 1988) representative target using a 12-parameter affine transformation. Movement correction and atlas transformation were accomplished in one resampling step (resulting in an isotropic 2 mm voxel size) to minimize blur and noise. Cross-modal image registration in patients was checked by comparing the optimized voxel similarity measure (or ETA score) to the 97.5 percentile obtained in the control group. When patients exhibited an ETA score below the threshold, even if for one transformation, we utilized anatomical image from a chronic (data not reported in this study) time point for the automatic normalization. In the case in which even the chronic time point exhibited an ETA score below the threshold, the anatomical images for the automatic normalization were selected after visual inspection by an expert author (AZS). N=9 cases out of 132 were inspected.

Lesion Segmentation

Lesions were manually segmented by students and post-doctoral fellows involved in the project using the Analyze biomedical imaging software system (Robb and Hanson, 1991). Wideman-one.com/gw/brain/analyze/formatdoc.htm) on the atlas-transformed T1-weighted MP-RAGE, T2-weighted spin echo images, and FLAIR images obtained at about 2 weeks post-stroke. Two board-certified neurologists (MC, AC) reviewed all segmentations, with special attention given to distinguishing lesion from CSF, hemorrhage from surrounding vasogenic edema, and identifying the number of lacunes and degree of periventricular white matter damage present. The edema in hemorrhagic strokes was included in the lesion. The periventricular white matter was rated according to the Longstreth et al. scale, where Grade 0 indicates no white matter abnormalities and Grade 9 indicates "very extensive and confluent" white matter disease (Longstreth et al., 1996). A neurologist (MC) reviewed all segmentations a second time paying special attention to the borders of the lesions, number of lacunes, and degree of white matter disease. The staff that was involved in segmenting or in reviewing the lesions was blind to the individual behavioral data. The lesions segmented in atlas space were summed at the voxel-wise level to display the number of patients with structural damage for each voxel at that location (lesion overlay map).

Lesion classification

The lesions were also automatically classified based on their overlap with three masks (gray matter, white matter subcortical regions including basal ganglia and thalamus) each computed as 50% conjunction of 38 single subject FreeSurfer (http://surfer.nmr.mgh.harvard.edu/) grey and white matter segmentations obtained from an independent group of healthy volunteers (age range=18-35) (Dale et al., 1999) on 1x1x1 mm MP-Rage T1-weighted images. Adequate segmentation was verified by inspection of the FreeSurfer-generated results in all subjects. The resulting grey matter

max consisted of 30,981 voxels. The white matter mask included all voxels not belonging to either cortical or subcortical gray matter mask above the brainstem. The mask computed in 1x1x1 mm space was re-sampled to 2x2x2 mm to match the lesion space. A K-means clustering in MatLab (MatLab Works) was run using the percent of lesion volume that intersected with each mask (i.e. number voxels in the lesion overlapping with each mask/total number of voxels in the lesion) as input, to display the overlap of each lesion group with gray matter, white matter, and subcortical nuclei.

Lesion-Behavior Machine Learning Methods

The voxel-wise relationship between stroke lesion maps and behavioral scores for each domain was analyzed using a multivariate leave-one-out ridge regression machine learning algorithm. We chose to use a linear ridge regression function to minimize bias but retain the ability to plot predictive weights back to brain anatomy (Phan et al., 2010). Prior to training the machine learning algorithm, transductive principal component analysis (PCA) (Zhu et al., 2008) was conducted on all subjects' segmented lesions. PCA was run on all 132 lesion binary voxel maps. The resulting components are maps that describe lesion variability. Components that explained 90% of variance in lesion location were retained, leaving 52 lesion components. Each subject's lesion could then be described using 52 weights. This enables a massive reduction of voxel data dimensionality while retaining meaningful differences in lesion distribution. Lesion weights and behavioral scores were then used to train a ridge regression algorithm that minimizes squared loss with a squared regularization term. The algorithm is as follows:

$$\arg\min_{\omega} \frac{1}{n}\sum_{i=1}^{n}(\vec{\omega}^T\vec{x}_i - \vec{y}_i)^2 + \lambda||\vec{\omega}||_2^2$$

The x vector indicates lesion location (in PCA space). The y vector contains behavioral factor scores for these same patients. The vector ω is the weight vector that describes the relative importance of each feature in x to the prediction of y. Lambda, a regularization coefficient, is determined empirically using a leave-one-out approach over a range of lambda values. We used the closed-form solution for our weight vector, which is as follows:

$$Let\ X = [\vec{x}_1, \vec{x}_2, \ldots \vec{x}_n]^T\ and\ Y = [\vec{y}_1, \vec{y}_2, \ldots \vec{y}_n]^T$$
$$\omega = (XX^T + \lambda I)^{-1}XY^T$$

In our leave-one-out approach, the closed-form solution was found using n-1 patients and used to estimate factor scores for the n^{th} patient. Regressor accuracy was determined using the square of the Pearson correlation coefficient between predicted and actual values, or % variance explained (r^2). The weight vector, ω, was then projected back in to native brain space to determine predictive voxels. It was additionally weighted by accuracy (r^2) so that the predictive power of voxels could be compared between behavioral domains. PCA and machine learning analyses was conducted in Matlab (the MathWorks Inc.). Ridge regression code was provided in part by Kilian

Weinberger, PhD, Associate Professor of Bioengineering at Washington University.

White matter pathways-behavior relationships

The relationship between white matter pathways damage and behavioral deficits was investigated using ridge regression. White matter pathway damage was assessed based on the intersection of the lesions with a probabilistic tractography atlas developed by (Thiebaut de Schotten et al., 2011) from 40 healthy right-handed adults in which reconstructed tracts were mapped within a common reference space (Montreal Neurological Institute). The tracts were reconstructed using spherical deconvolution, a method that allows for the estimation of multiple orientations in voxels containing different populations of crossing fibers (Anderson, 2005; Tournier et al., 2004). The atlas includes 57 tracts each at 20, 40, 60, 80, and 100% probabilities of overlap across healthy subjects. The 57 tracts were thresholded at different probabilities and then overlaid in a single image to measure tract overlap with the lesion-behavior map for different domains. Ridge regression was used to examine whether there was any significant relationship between number of behavioral domain affected and number of white matter tracts.

Anderson, A.W. (2005). Measurement of fiber orientation distributions using high angular resolution diffusion imaging. Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine 54, 1194-1206.
Benedict, R.H.B. (1997). Professional manual of the Brief Visuospatial Memory Test-revised. (Psychological Assesment Resources.).
Brandt, J., and Benedict, R.H.B. (2001). Professional manual of the Hopkins Verbal Learning Test-revised. (Psychological Assesment Resources.).
Connor, L.T., DeShazo Braby, T., Snyder, A.Z., Lewis, C., Blasi, V., and Corbetta, M. (2006). Cerebellar activity switches hemispheres with cerebral recovery in aphasia. Neuropsychologia 44, 171-177.
Dale, A.M., Fischl, B., and Sereno, M.I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage 9, 179-194.
Delis, D.C., Kaplan, Edith, Kramer, Joel H. (2001). Delis-Kaplan Executive Function System (D-KEFS) (San Antonio, TX: The Psychological Corporation).
Demeurisse, G., Demol, O., and Robaye, E. (1980). Motor evaluation in vascular hemiplegia. European neurology 19, 382-389.
Dreeben-Irimia, O. (2008). Physical Therapy Clinical Handbook for PTAs, 1 edn (Jones & Bartlett Learning).
Goodglass, H., Kaplan, E., and Barresi, B. (2001). Boston Diagnostic Aphasia Examination., 3 edn (Austin, TX.: Pro-Ed).
Hachinski, V., Iadecola, C., Petersen, R.C., Breteler, M.M., Nyenhuis, D.L., Black, S.E., Powers, W.J., DeCarli, C., Merino, J.G., Kalaria, R.N.*, et al.* (2006). National Institute of Neurological Disorders and Stroke-Canadian Stroke Network vascular cognitive impairment harmonization standards. Stroke 37, 2220-2241.
Keith, R.A., Granger, C.V., Hamilton, B.B., and Sherwin, F.S. (1987). The functional independence measure: a new tool for rehabilitation. Adv Clin Rehabil 1, 6-18.
Kempen, J.C., de Groot, V., Knol, D.L., Polman, C.H., Lankhorst, G.J., and Beckerman, H. (2011). Community walking can be assessed using a 10-metre timed walk test. Mult Scler 17, 980-990.
Longstreth, W.T., Jr., Manolio, T.A., Arnold, A., Burke, G.L., Bryan, N., Jungreis, C.A., Enright, P.L., O'Leary, D., and Fried, L. (1996). Clinical correlates of white matter findings on cranial magnetic resonance imaging of 3301 elderly people. The Cardiovascular Health Study. Stroke 27, 1274-1282.
Mesulam, M.-M., MD (1985). Patterns In Behavioral Neuroanatomy: Association Areas, The Limbic System and Hemispheric Specialization. In Principles of Behavioral Neurology, M.-M. Mesulam, MD, ed. (Philadelphia: F.A. Davis), pp. 1-70.
Oxford Grice, K., Vogel, K.A., Le, V., Mitchell, A., Muniz, S., and Vollmer, M.A. (2003). Adult norms for a commercially available Nine Hole Peg Test for finger dexterity. Am J Occup Ther 57, 570-573.
Perry, J., Garrett, M., Gronley, J.K., and Mulroy, S.J. (1995). Classification of walking handicap in the stroke population. Stroke 26, 982-989.
Posner, M.I., Walker, J.A., Friedrich, F.J., and Rafal, R.D. (1984). Effects of parietal injury on covert orienting of attention. Journal of Neuroscience 4, 1863-1874.

Rengachary, J., d'Avossa, G., Sapir, A., Shulman, G.L., and Corbetta, M. (2009). Is the posner reaction time test more accurate than clinical tests in detecting left neglect in acute and chronic stroke? Archives of physical medicine and rehabilitation 90, 2081-2088.
Robb, R.A., and Hanson, D.P. (1991). A software system for interactive and quantitative visualization of multidimensional biomedical images. Australas Phys Eng Sci Med 14, 9-30.
Rorden, C., and Karnath, H.O. (2010). A simple measure of neglect severity. Neuropsychologia 48, 2758-2763.
Rowland, D.J., Garbow, J.R., Laforest, R., and Snyder, A.Z. (2005). Registration of [18F]FDG microPET and small-animal MRI. Nucl Med Biol 32, 567-572.
Talairach, J., and Tournoux, P. (1988). Co-Planar Stereotaxic Atlas of the Human Brain (New York: Thieme Medical Publishers, Inc.).
Thiebaut de Schotten, M., Ffytche, D.H., Bizzi, A., Dell'Acqua, F., Allin, M., Walshe, M., Murray, R., Williams, S.C., Murphy, D.G., and Catani, M. (2011). Atlasing location, asymmetry and inter-subject variability of white matter tracts in the human brain with MR diffusion tractography. Neuroimage 54, 49-59.
Tombaugh, T.N., Kozak, J., and Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. Arch Clin Neuropsychol 14, 167-177.
Tournier, J.D., Calamante, F., Gadian, D.G., and Connelly, A. (2004). Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. Neuroimage 23, 1176-1185.
van der Lee, J.H., Beckerman, H., Lankhorst, G.J., and Bouter, L.M. (2001). The responsiveness of the Action Research Arm test and the Fugl-Meyer Assessment scale in chronic stroke patients. J Rehabil Med 33, 110-113.
Wilson, B., Cockburn, J., and Halligan, P. (1987). Development of a behavioral test of visuospatial neglect. Archives of physical medicine and rehabilitation 68, 98-102.
Zhu, Y., Wu, Y., Liu, X., and Mio, W. (2008). Transductive optimal component analysis. 19th International Conference on Pattern Recognition 2008, 1–4.

Domain	Function tested	Test	Score Recorded
Motor	Range of Motion	AROM: Shoulder flexion	Left Shoulder flexion
			Right Shoulder flexion
			Left Wrist Extension
			Right Wrist Extension
	Strength	Jamar Dynamometer	Left Grip strength
			Right Grip Strength
	Dexterity	Nine-hole peg test	Left hand pegs/second
			Right hand pegs/second
	Dexterity and Range of motion	ARAT	Left total
			Right total
	Walking	Timed Walk	Index of Timed Walk + FIM Walk Item
		FIM Walk Item	
	Strength	Motricity Index	Left Lower Extremity total
			Right Lower Extremity total
	Range of Motion	AROM: Lower extremity	Left Ankle dorsiflexion
			Right Ankle dorsiflexion
Language	Comprehension	BDAE: Comprehension	Basic Word Discrimination
			Commands
			Complex Ideational Material
	production, semantic	BDAE: Expression	Boston Naming Short Form
	comprehension	BDAE: Reading	Oral Reading of Sentences
			Comprehension of Oral Reading of Sentences
	production, phonological	Experimental measures	Nonword Reading
			Stem Completion
	production, semantic	Verbal Fluency	Animal Naming test
Memory	Spatial, recall	BVMT	Immediate Total Recall T-score
			Delayed Recall T-score
			Delayed Recall percent retained
	Spatial, recognition		Delayed Recognition Discrimination Index
	Verbal, recall	HVLT	Immediate Total Recall T-score
			Delayed Recall T-score
			Delayed Recall percent retained
	Verbal, recognition		Delayed Recognition Discrimination Index
Visual Attention	Spatial, recall	Spatial Span	Span Forward
			Span Backwards
	Visual field	Posner orienting task, reaction time	Visual field effect [Left-Right], RT
	Shifting		Validity effect [Valid-Invalid], RT
	Average		Overall performance, RT
	Shifting		Disengagement effect [(LI-LV)-(RI-RV)], RT
	Visual field	Posner orienting task, accuracy	Visual field effect [Left-Right], accuracy
	Shifting		Validity effect [Valid-Invalid], accuracy
	Average		Overall performance, accuracy
	Shifting		Disengagement effect [(LI-LV)-(RI-RV)], accuracy
	Visual Field	BIT star cancellation	Center-of-cancellation, Left vs. Right
		Mesulam unstructured symbol cancellation	Center-of-cancellation, Left vs. Right

Supplementary Table 1 (refers to Supplementary Methods)**.**

Domains of function, function tested, list of tests, and score recorded.

	Source population	**Study sample**	Healthy controls
Age	chi-square asymptotic sig.<0.01 for source population vs. study sample		
18-30	3%	**4%**	3%
31-50	20% *	**30%**	23%
51-70	47% **	**63%**	68%
71 or older	31% **	**4%**	6%
Gender	No significant differences among groups		
Female	48%	**45%**	52%
Male	52%	**55%**	48%
Race	chi-square asymptotic sig.<0.01 for source population vs. study sample		
Caucasian	62%**	**34%**	32%
African- American	35%**	**64%**	61%
Other	2%	**2%**	6%
Education	No significant differences among groups		
Incomplete high school	22%	**16%**	16%
High school	40%	**39%**	29%
Incomplete college	23%	**25%**	35%
College	8%	**8%**	10%
Post-graduate	7%	**11%**	10%
Predisposing Factors	*Binomial test p<0.05 for starred items		
Hypertension	73%	**70%**	26%**
Diabetes Mellitus	29%	**31%**	16%**
Coronary Artery Disease	22%**	**8%**	6%
Atrial fibrillation	11%*	**5%**	3%
Depression	11%*	**5%**	3%

*p<0.01; **p<0.001

Supplementary Table 2 (refers to result section 'Clinical representativeness of study sample')**.** Statistical comparisons of demographic and risk factors for stroke.

Language battery	Variance explained: 76.8%	
Component Matrix		
	Component	
	1	2
Word Comprehension	.787	
Commands	.799	
Complex Ideational Material	.984	
Boston Naming Test	.711	.320
Oral Reading of Sentences	.523	.477
Comprehension of Oral Reading of Sentences	.558	.378
Non-word Reading		.955
Stem completion	.444	.599
Animal Naming	.871	

Supplementary Table 3 (refers to result section 'Behavior: within domain factor analysis'). Principal component analysis of language battery on patients with aphasia, i.e. with language score >2SD from controls.

Higher-order PCA	Variance explained: 69%		
Rotated Component Matrix			
	Component		
	1	2	3
Language	.881		
Memory: Verbal	.890		
Memory: Spatial	.594	.531	
Motor: Left Limb		.772	
Motor: Right Limb			.853
Attention: Visual Field effect		.799	-.351
Attention: Validity / Disengagement effect			.676
Attention: Average performance		-.662	

Supplementary Table 4 (refers to Figure 3). Higher-order principal component across domains and loading for individual factor scores.

All variables	Variance explained: 47.8%		
Rotated Component Matrix			
	Component		
	1	2	3
Basic Word Discrimination	.421		
Commands	.575		
Complex Ideational Material	.778		
Boston Naming Short Form	.670		
Oral Reading of Sentences	.665		
Comprehension of Oral Reading of Sentences	.629		
Nonword Reading	.670		
Stem Completion	.635		
Animal Naming	.607		
Left Shoulder flexion		.895	
Right Shoulder flexion			.877
Left Wrist Extension		.819	
Right Wrist Extension			.833
Left Grip strength		.587	
Right Grip strength			.667
Left hand pegs/second Nine-hole peg test		.703	
Right hand pegs/second Nine-hole peg test			.784
Left total ARAT		.885	
Right total ARAT			.911
Timed Walk + FIM Walk Item		.529	.491
Left Lower Extremity total Motricity Index		.885	
Right Lower Extremity total Motricity Index			.844
Left Ankle dorsiflexion AROM		.899	
Right Ankle dorsiflexion AROM			.819
Posner Visual field effect L-R, RT		-.452	.301
Posner Visual field effect L-R, accuracy		-.649	
Posner Validity effect V-I, RT			
Posner Validity effect V-I, accuracy			
Posner Disengagement effect [(LI-LV)-(RI-RV)], RT			
Posner Disengagement effect [(LI-LV)-(RI-RV)], accuracy			
Posner Overall performance, RT	-.492		
Posner Overall performance, accuracy	.425	.558	
Mesulam Center-of-cancellation, L-R misses		-.620	
BIT: star cancellation, Center-of-cancellation, L-R misses			
BVMT Immediate Total Recall T-score	.669		
BVMT Delayed Recall T-score	.679		
BVMT Delayed Recall percent retained			
BVMT Delayed Recognition discrimination Index	.545		
HVLT Immediate Total Recall T-score	.730		
HVLT Delayed Recall T-score	.735		
HVLT Delayed Recall percent retained	.565		
HVLT Delayed Recognition discrimination Index	.680		
Spatial Span Span Forward	.654		
Spatial Span Span Backwards	.623	.374	

Supplementary Table 5 (refers to Figure 4)**.** PCA on raw scores of neuropsychological tests.

ID	Lesion Volume (in voxels of 2x2x2mm)	Lesion Etiology	Periventricular white matter disease rating	Lacunae	Cortical GMmask	Subcortical GMmask	White matter mask	Expert localization
24	4233	Ischemic	1	0	19%	0%	80%	cortico-subcortical
26	4563	Ischemic	1	1	0%	0%	0%	cerebellar
27	7574	Ischemic	2	2	23%	26%	48%	cortico-subcortical
29	776	Ischemic	2	0	75%	0%	23%	cortical
30	5332	Ischemic	1	2	71%	0%	23%	cortical
32	794	Ischemic	1	0	0%	0%	0%	cerebellar
33	910	Ischemic	1	2	52%	0%	44%	cortical
35	2462	Ischemic	1	0	66%	0%	28%	cortical
36	6907	Ischemic	1	0	40%	14%	39%	cortico-subcortical
37	397	Ischemic	2	2	0%	95%	5%	subcortical
38	390	Ischemic	2	0	0%	16%	78%	subcortical
39	6832	Ischemic	0	0	59%	1%	38%	cortico-subcortical
40	5997	Ischemic	1	0	44%	17%	33%	cortico-subcortical
41	4161	Ischemic	2	1	30%	0%	70%	cortical
43	4275	Ischemic	1	0	2%	30%	62%	subcortical
44	420	Ischemic	1	2	0%	1%	0%	brainstem
45	126	Ischemic	0	0	0%	0%	100%	white matter only
47	455	Ischemic	1	0	6%	11%	82%	subcortical
48	2629	Ischemic	1	0	3%	30%	64%	subcortical
49	303	Ischemic	1	0	0%	0%	0%	brainstem
50	3377	Ischemic	1	0	0%	0%	0%	cerebellar
51	104	Ischemic	1	2	0%	9%	90%	subcortical
53	3255	Hemorrhagic	0	1	27%	6%	66%	cortico-subcortical
55	374	Ischemic	0	2	0%	14%	86%	subcortical
56	1399	Ischemic	0	3	33%	0%	66%	cortical
58	294	Ischemic	1	0	0%	13%	86%	subcortical
60	371	Ischemic	2	0	0%	11%	0%	brainstem

62	7738	Ischemic	0	2	54%	2%	35%	cortical
63	5177	Ischemic	0	2	74%	0%	23%	cortical
64	1164	Ischemic	0	1	12%	1%	81%	cortical
65	119	Ischemic	0	0	0%	0%	0%	brainstem
67	118	Ischemic	5	2	0%	0%	100%	white matter only
68	9760	Hemorrhagic	3	2	27%	17%	54%	cortico-subcortical
69	737	Ischemic	1	0	0%	83%	8%	subcortical
71	13	Ischemic	0	2	0%	0%	0%	brainstem
72	10186	Ischemic	2	5	57%	0%	41%	cortical
73	2275	Ischemic	0	0	65%	0%	17%	cortical
74	146	Ischemic	0	0	62%	0%	36%	cortical
75	154	Ischemic	0	2	1%	0%	99%	white matter only
76	8006	Ischemic	0	1	25%	6%	68%	cortico-subcortical
77	331	Ischemic	0	2	0%	0%	0%	brainstem
78	991	Ischemic	2	6	18%	0%	82%	cortical
79	149	Ischemic	0	0	0%	6%	92%	subcortical
80	1332	Ischemic	0	2	0%	0%	0%	cerebellar
81	111	Ischemic	0	2	0%	64%	36%	subcortical
82	5906	Ischemic	0	0	76%	0%	21%	cortical
83	336	Ischemic	2	1	51%	0%	48%	cortical
84	323	Ischemic	0	2	66%	0%	34%	cortical
85	584	Ischemic	0	0	70%	0%	29%	cortical
87	2565	Ischemic	0	0	60%	0%	37%	cortical
88	2877	Hemorrhagic	5	15	7%	33%	45%	subcortical
90	91	Ischemic	0	1	0%	0%	0%	brainstem
92	470	Ischemic	0	0	50%	0%	37%	cortical
93	1003	Ischemic	4	0	3%	21%	76%	subcortical
95	10110	Ischemic	0	0	24%	26%	44%	other
97	771	Ischemic	0	0	42%	0%	56%	cortical
98	233	Hemorrhagic	2	2	0%	9%	0%	brainstem
99	13895	Ischemic	0	0	57%	5%	28%	cortico-subcortical
100	5160	Ischemic	0	2	51%	1%	46%	cortico-subcortical
101	2007	Ischemic	0	2	5%	0%	0%	cerebellar

102	162	Ischemic	0	0	0%	99%	0%	subcortical
103	11301	Hemorrhagic	0	0	26%	21%	50%	cortico-subcortical
104	162	Ischemic	6	2	7%	1%	92%	white matter only
105	4020	Hemorrhagic	0	0	54%	1%	43%	cortico-subcortical
106	427	Ischemic	1	1	8%	0%	92%	white matter only
107	2637	Hemorrhagic	1	1	2%	34%	43%	subcortical
108	6287	Ischemic	1	1	69%	0%	23%	cortical
109	14768	Ischemic	1	1	59%	3%	34%	cortico-subcortical
110	24607	Ischemic	1	1	57%	2%	40%	cortico-subcortical
111	1718	Ischemic	1	2	4%	0%	96%	white matter only
112	1224	Ischemic	5	1	2%	20%	78%	subcortical
114	7920	Ischemic	1	1	53%	7%	33%	cortical
115	7104	Ischemic	1	15	73%	0%	16%	cortical
116	313	Ischemic	1	1	22%	0%	76%	cortical
117	8188	Hemorrhagic	1	1	24%	17%	57%	cortico-subcortical
118	8827	Ischemic	0	2	44%	0%	15%	other
119	3911	Ischemic	1	1	64%	0%	30%	cortical
120	2833	Ischemic	1	1	47%	2%	50%	cortical
122	6975	Ischemic	1	1	81%	0%	14%	cortical
123	237	Ischemic	1	3	0%	13%	87%	subcortical
124	324	Ischemic	1	2	58%	0%	26%	cortical
125	52	Ischemic	4	1	0%	0%	0%	brainstem
126	3077	Hemorrhagic	1	0	2%	60%	29%	subcortical
128	34627	Hemorrhagic	0	1	51%	8%	34%	cortico-subcortical
129	27901	Ischemic	0	0	56%	6%	34%	cortico-subcortical
133	9475	Ischemic	1	0	76%	0%	16%	cortical
135	5369	Ischemic	0	0	11%	38%	45%	cortico-subcortical
136	362	Ischemic	0	1	82%	0%	18%	cortical
138	2782	Ischemic	0	2	3%	37%	50%	subcortical
140	10300	Ischemic	0	0	27%	21%	49%	cortico-subcortical
142	400	Ischemic	2	0	0%	47%	51%	subcortical
143	999	Ischemic	0	0	1%	17%	80%	subcortical
144	2714	Ischemic	0	0	59%	2%	38%	cortical

145	192	Ischemic	0	0	3%	45%	52%	subcortical
150	6103	Ischemic	2	0	19%	23%	55%	cortico-subcortical
151	1451	Ischemic	0	0	12%	0%	88%	white matter only
152	8412	Ischemic	3	2	56%	0%	34%	cortico-subcortical
154	201	Ischemic	0	0	1%	18%	78%	subcortical
155	732	Ischemic	1	0	0%	0%	0%	cerebellar
157	3262	Hemorrhagic	5	5	3%	55%	32%	subcortical
158	12208	Hemorrhagic	1	1	23%	16%	57%	cortico-subcortical
160	8465	Hemorrhagic	1	4	26%	16%	54%	cortico-subcortical
161	8397	Ischemic	0	0	15%	31%	47%	cortico-subcortical
162	5285	Ischemic	0	0	36%	0%	61%	cortico-subcortical
163	1222	Ischemic	4	5	33%	5%	46%	cortical
164	3650	Ischemic	0	0	52%	0%	40%	cortical
165	255	Ischemic	0	5	67%	0%	24%	cortical
166	2657	Ischemic	6	3	55%	0%	39%	cortical
167	6315	Ischemic	0	1	59%	1%	38%	cortico-subcortical
168	179	Ischemic	5	10	41%	0%	56%	cortical
169	596	Ischemic	5	4	3%	20%	77%	subcortical
170	6182	Hemorrhagic	0	2	13%	33%	46%	other
171	661	Ischemic	2	0	2%	0%	0%	cerebellar
172	4744	Ischemic	0	1	31%	15%	52%	cortico-subcortical
173	2711	Hemorrhagic	2	5	0%	67%	15%	subcortical
174	26935	Ischemic	0	0	58%	4%	30%	cortico-subcortical
175	538	Ischemic	0	0	0%	0%	0%	cerebellar
178	5116	Hemorrhagic	2	0	15%	31%	51%	subcortical
179	582	Ischemic	2	0	0%	58%	42%	subcortical
180	3332	Ischemic	0	2	16%	0%	84%	cortico-subcortical
181	1804	Hemorrhagic	6	4	0%	1%	0%	brainstem
182	660	Hemorrhagic	2	18	3%	71%	21%	subcortical
183	1363	Ischemic	1	0	51%	0%	47%	cortical
186	70	Ischemic	3	6	0%	0%	0%	brainstem
187	5521	Hemorrhagic	3	2	50%	0%	45%	cortico-subcortical
188	5005	Hemorrhagic	4	0	3%	46%	33%	subcortical

190	2625	Hemorrhagic	5	2	51%	0%	48%	cortico-subcortical
192	13627	Ischemic	2	2	65%	0%	33%	cortico-subcortical
193	126	Ischemic	2	4	0%	0%	0%	brainstem
194	19211	Ischemic	1	0	37%	17%	49%	cortico-subcortical
195	17134	Hemorrhagic	2	0	47%	6%	44%	cortico-subcortical
196	867	Ischemic	1	2	59%	0%	21%	cortical
Average	**4272**		**1.2**	**1.6**	**28%**	**12%**	**41%**	
Percentage		**Hemorragic=17%**						
		Ischemic=83%						

Supplementary table 6 (refers to Figure 5). Individual lesion information with etiology, degree of periventricular white matter disease; number of lacunes, and percentage of subcortical, cortical, and white matter damage.

LESION TOPOGRAPHY	Corbetta 2015 Automatic clustering (Total N=132)	Kang 2003 Visual inspection (Total n=172; Single lesions n=104)	Wessel 2006 Visual inspection (Total n=510; Single lesions n=302)
Cortical	13	16	14
Cortico-subcortical	23	33	16
Subcortical (basal ganglia+thalamus)	16	50	66
White matter only	23		
Brainstem	6		
Cerebellum	17		

Supplementary table 7. Lesion topography (refers to Figure 5)**.** The relative frequency (% of total number) of lesion location in different prospective series. In our study localization by automatic clustering based on overlap with gray/white matter masks. In other studies localization by visual inspection.

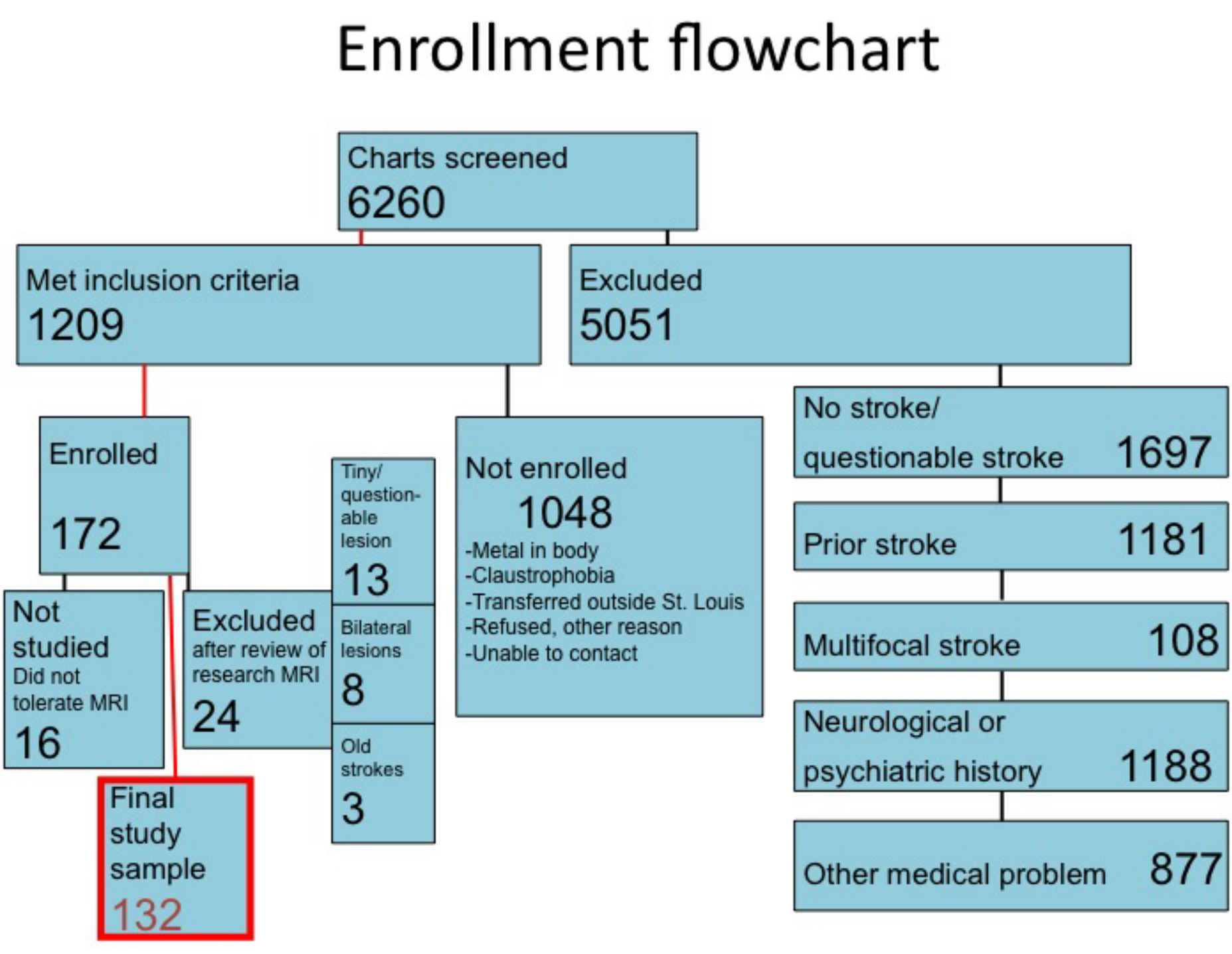
Lyden 1999

NIHSS Item	Factor 1	Factor 2	Factor 3	Factor 4	
Language	0.97				Left hemisphere
LOC questions	0.84				
LOC commands	0.77				
Right arm			0.98		
Right leg			0.89		
Dysarthria			0.49		
Limb Ataxia	excluded from analysis				
Gaze		0.71			Right hemisphere
Visual		0.69			
Extinction/neglect		0.69			
Sensory		0.60			
Left arm				1	
Left leg				0.87	
Facial Palsy	excluded from analysis				

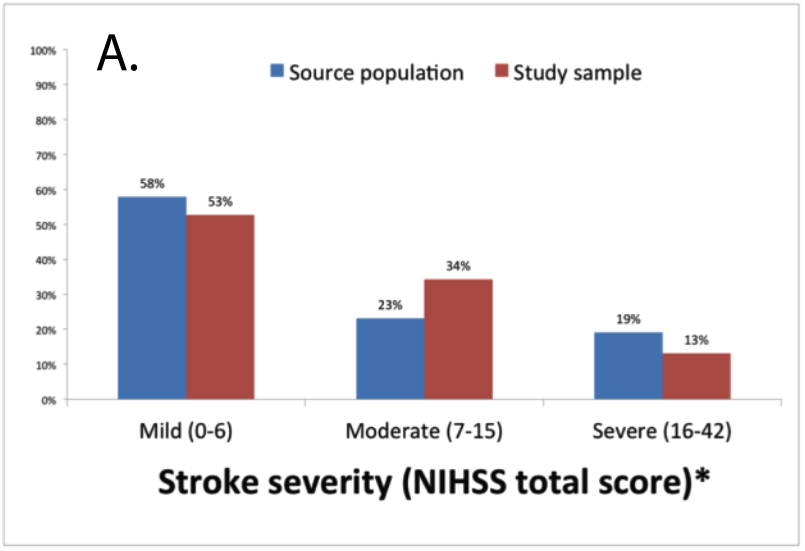
Corbetta 2015

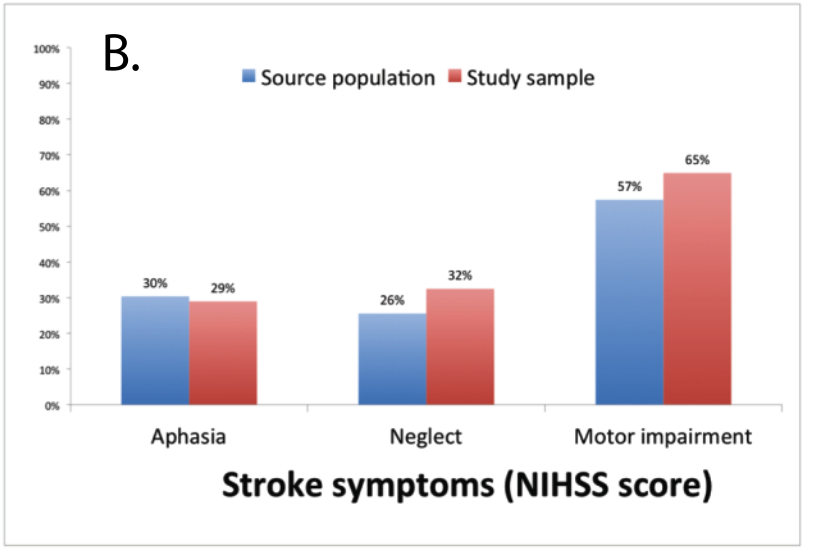
NIHSS Item	Factor 1	Factor 2	Factor 3	Factor 4	
Language			.810		Left hemisphere
LOC questions			0.77		
LOC commands			0.87		
Right arm	0.93				
Right leg	0.90				
Dysarthria	0.56				
Limb Ataxia				0.48	
Gaze				0.88	Right hemisphere
Visual				0.68	
Extinction/neglect		0.75			
Sensory		0.38			
Left arm		0.91			
Left leg		0.92			
Facial Palsy	0.42	0.59			

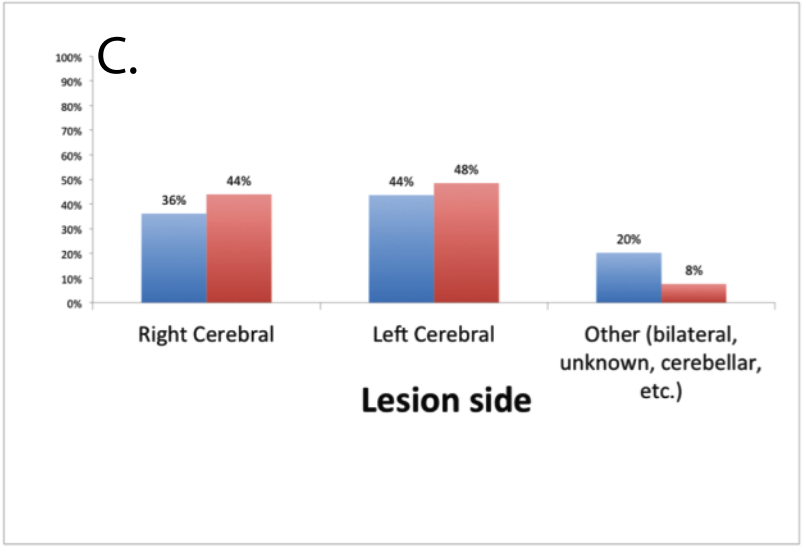
Supplementary table 8. Principal component analysis of NIHSS scores across two different studies (refers to Discussion). Note that both in Lyden 1999 and this study a few components (one motor and one cognitive for each hemisphere) account for the majority of variance.

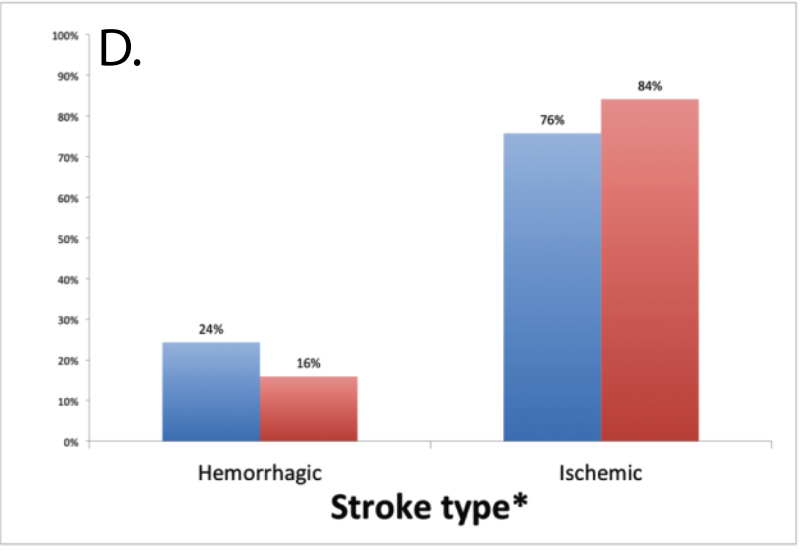


Supplementary Figure 1 (refers to Methods and Supplementary information)**.** Enrollment flowchart.

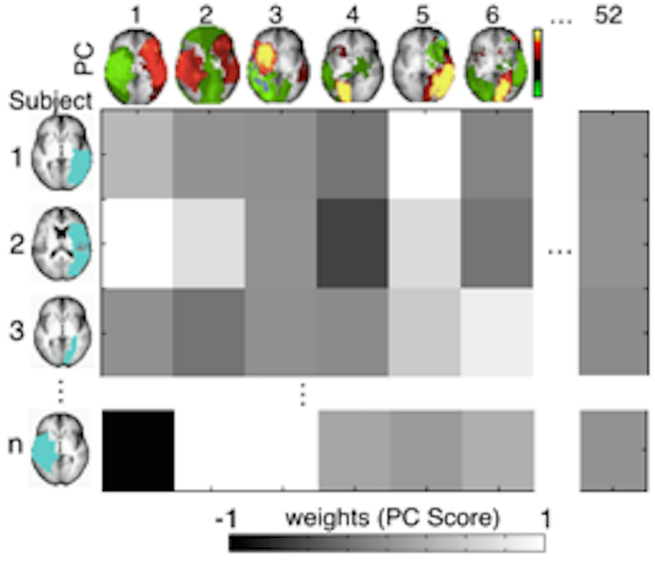








Supplementary Figure 2 (refers to Result section 'Clinical representativeness of study sample')**.** Comparison between study sample (n=132) and source population (n=1209) in terms of neurological variables: A) NIH Stroke Scale severity; B) Frequency of neurological deficits; C) Side of stroke; D) Type of stroke.

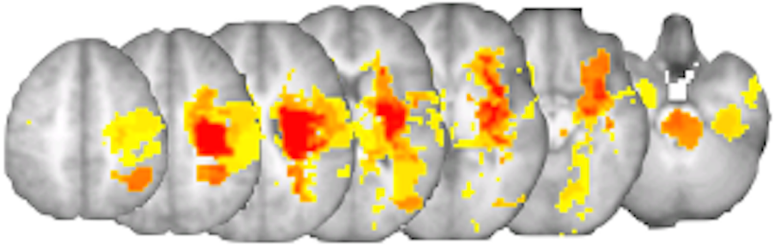


Ridge Regression

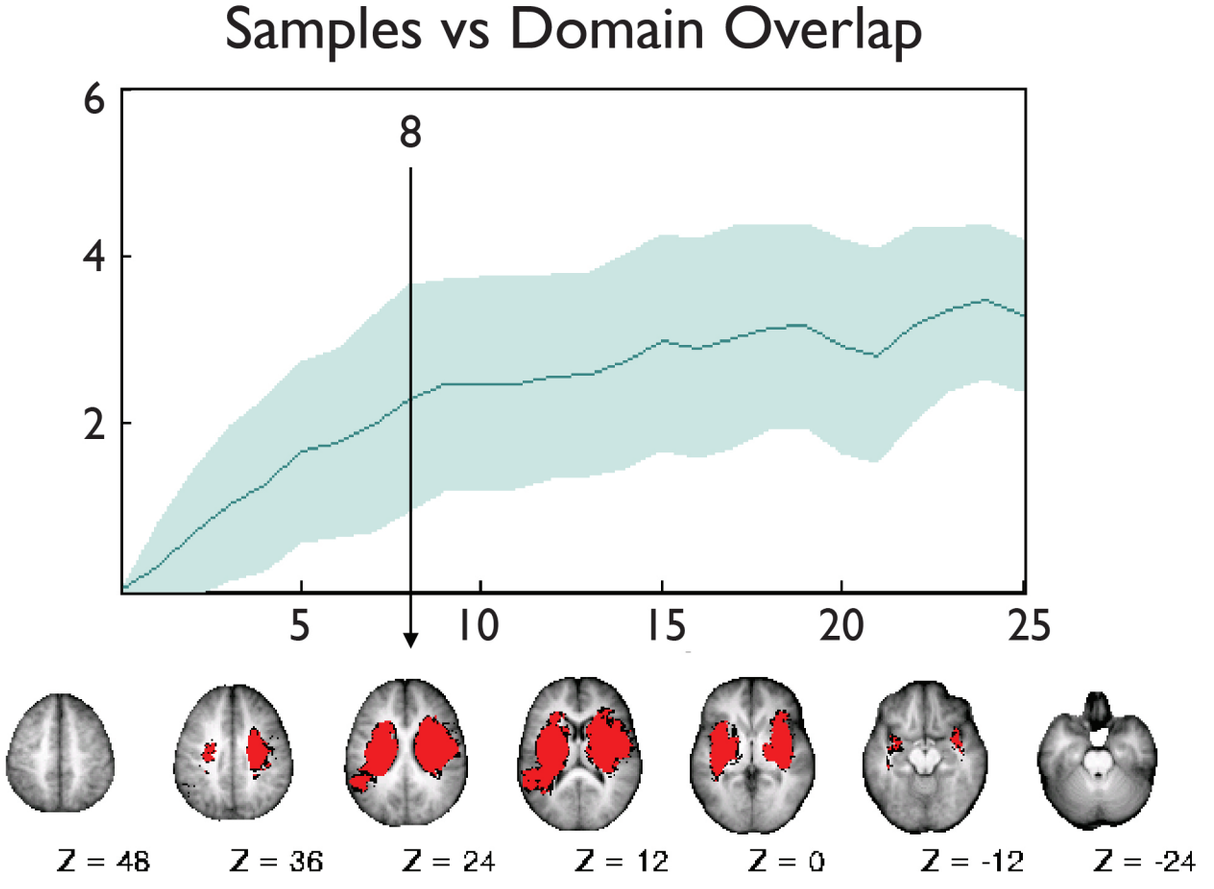
$$= \vec{x} \rightarrow \sum_{i=1}^{n} (\vec{\omega}^T \vec{x}_i - \vec{y}_i)^2 + \lambda ||\vec{\omega}||_2^2$$

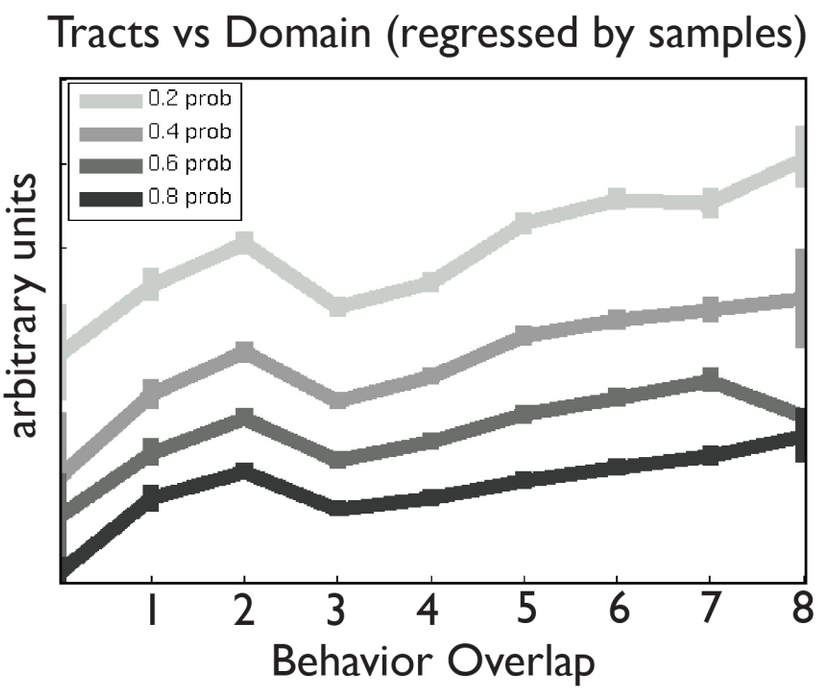
	Predict $(\vec{\omega}^T \vec{x}_n)$	Actual $(\vec{y}_n)$
L motor	-2.6	-3
R motor	0.7	1
Language	0.1	0.4
Memory	0.3	-0.7
Attention	-1.7	-2.2

Left Motor Weight (ω) Map



Supplementary Figure 3. Ridge regression (refers to result section 'Behavior-to-Anatomy relationships: single domain prediction'**).**
A schematic of applying ridge regression to use lesion to predict deficit in each domain and produce lesion-symptom maps for the group.
Left: voxel lesion maps shown in blue. PCA on all lesions was used to generate lesion components. Lesions are projected from 65k voxels to 52 components.
Then, a ridge regression function using lesion components (x) to explain deficit (y) is trained for n-1 subjects.
This function can then generate a prediction of deficit in each domain based on data from patient n.
Additionally, the weight matrix ω solved for a given behavioral domain can be projected back on to the brain.
Right: the weights used to predict left motor deficit. This was same approach was taken to produce lesion-symptom maps for all behavioral domains.

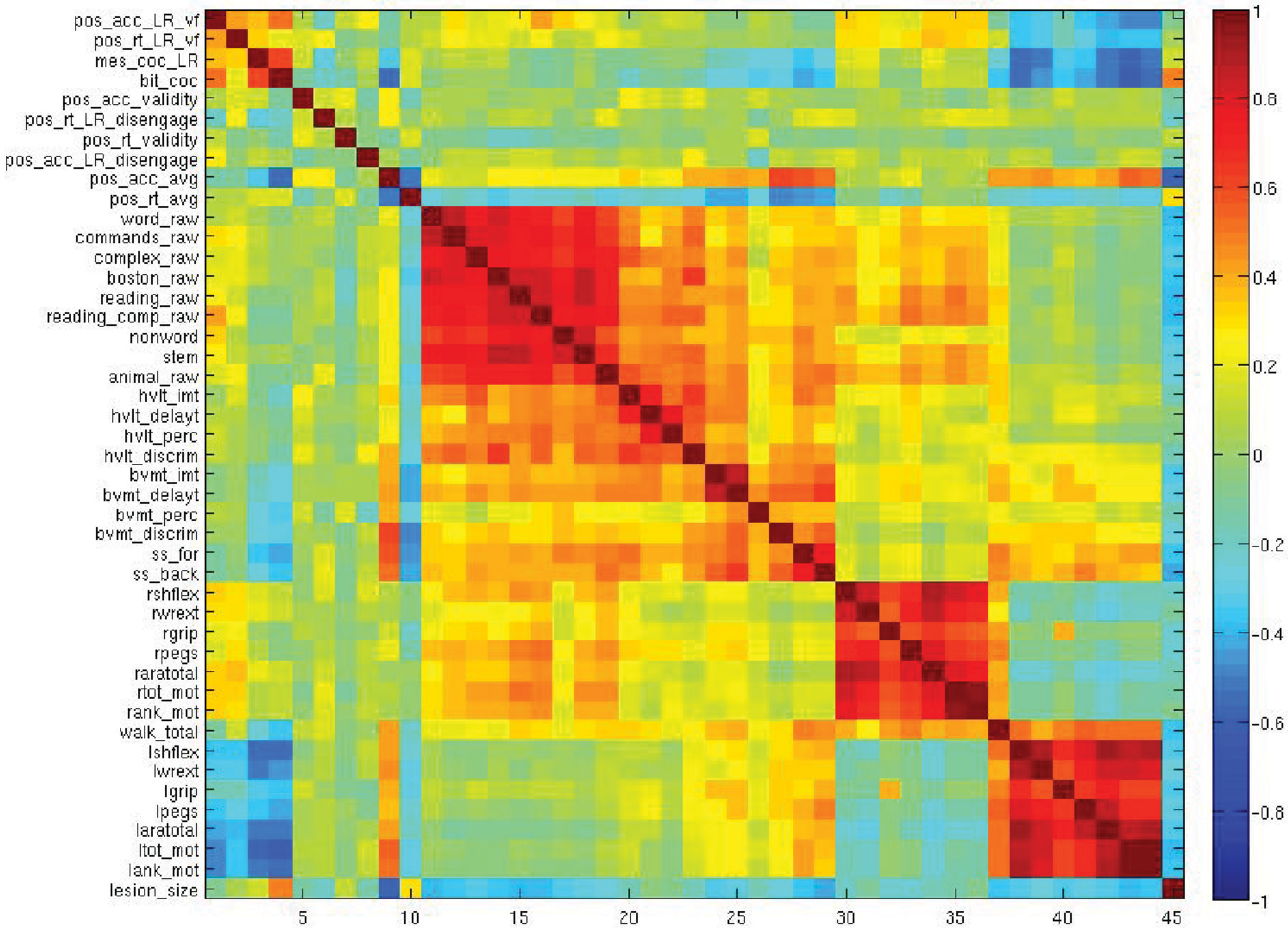




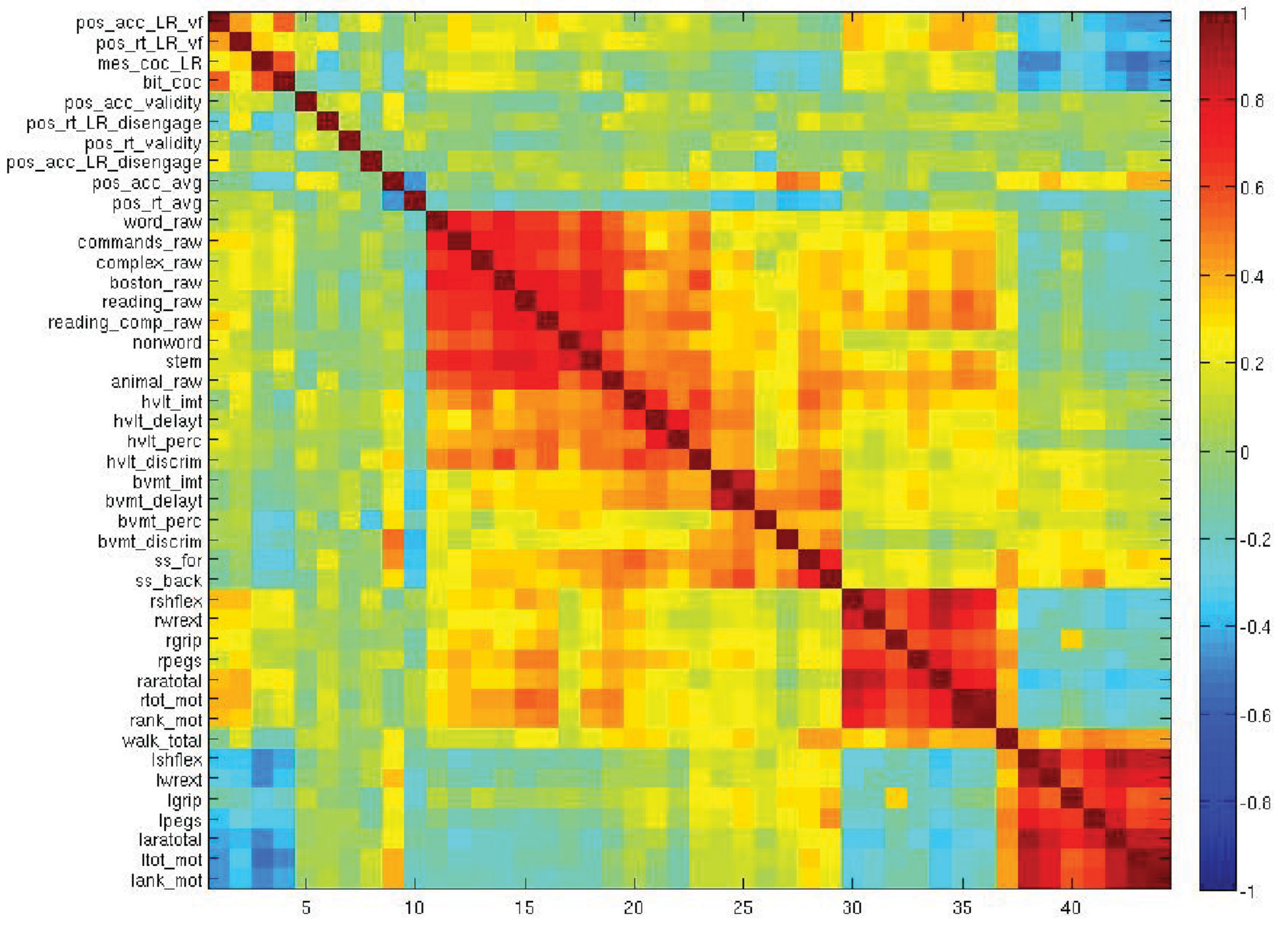
probability	0.2	0.4	0.6	0.8
rho	0.147	0.155	0.154	0.125
p-value	<0.0001	<0.0001	<0.0001	<0.0001

Supplementary Figure 4. Lesion sample at each voxel vs. # behavioral deficit (refers to Fig.7)**.** Correlation between domain overlap and white matter tract overlap is independent on lesion samples. Left: in voxels with less than 8 lesion subjects, a relationship between number of samples and domain conjunction was observed. To control for this, only voxels in which at least 8 subjects had lesion were used (bottom left).
Right: Domain overlap is plotted against tract overlap. A one-tailed t-test performed at all four tract probabilities demonstrates that voxels associated with 3+ domains contained more tracts on average than voxels associated with <3 domains (bottom right).

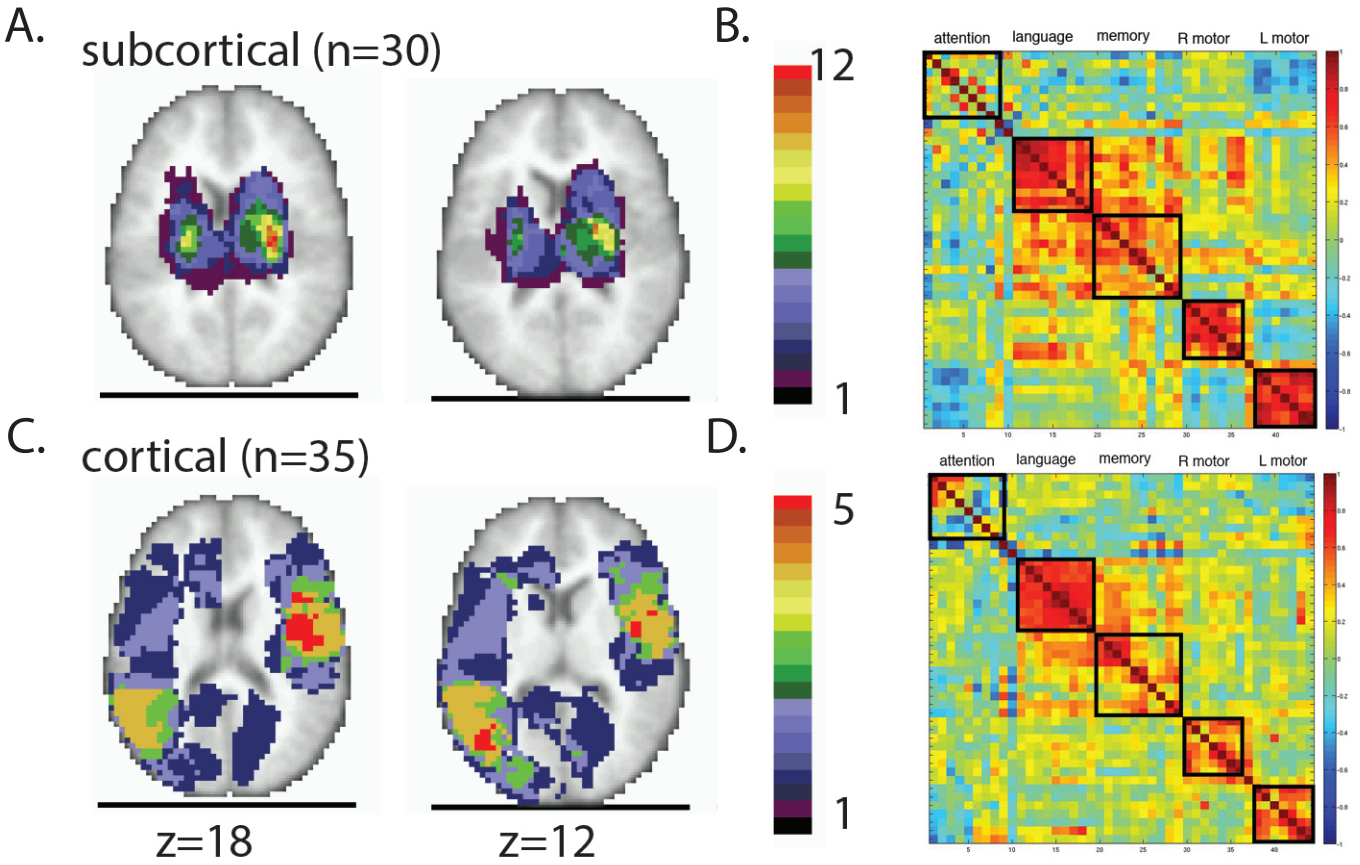
A.



B.



Supplementary Figure 5. Correlation matrix of behavioral scores: lesion volume. (refers to result section 'Control for lesion volume and location'). A). Total correlation; B). Partial correlation with lesion volume removed. Note similarity of correlation matrices without (A) or with (B) lesion volume removed.



Supplementary Figure 6. Correlation matrix of behavioral scores: subcortical vs. cortical lesions (refers to result section 'control for lesion volume and location'). A) Subcortical strokes (thalamus and basal ganglia); B) Correlation matrix of behavioral scores for subcortical group; C) Cortical strokes heterogeneous lesion location; D) Correlation matrix of behavioral scores for cortical group. Note similarity of correlation matrices despite variability in lesion location for the two groups.The spatial correlation between the matrices is 0.63.